Institution:	**University of Oxford**
Unit of Assessment:	**11 Computer Science and Informatics**
Title of case study:	**Securing Data with Database Firewall (3)**

1. Summary of the impact (indicative maximum 100 words)

Pioneering research into Inductive Logic Programming in the UOA led to the creation of Secerno Ltd. From 2008 Secerno attracted investment of approximately $20m and successfully released several updated versions of its product DataWall, based on this Oxford research. In May 2010 Oracle Corporation bought Secerno specifically to gain access to this technology, which now forms a core part of Oracle's database protection and compliance products. Oracle continues to develop the software, which is used across the globe by public entities and private companies to protect databases from internal and external attack and to ensure that they comply with relevant legislation. Customers include major businesses such as T-Mobile, which uses Database Firewall to protect 35 million users.

2. Underpinning research (indicative maximum 500 words)

Inductive Logic Programming (ILP) is a symbolic machine learning technique which helps extend scientific theories by using automated systems to generate new rules, from data and existing theories. The technique was established by Stephen Muggleton whilst he was in Oxford University's Programming Research Group (PRG) as an EPSRC Fellow (1993-1997); a 1994 paper outlined the key aspects of the new field, including a 'model theory' for ILP, a generic ILP algorithm and a 'proof theory', as well as applications and implementations of ILP [1]. ILP was latterly implemented as *Progol* in 1995 [2]. ILP was extended by probability distributions as *Stochastic Logic Programming* [3]. Progol was used in building parts-of-speech taggers from very large example corpuses in natural language applications [4]. Ashwin Srinivasan developed an ILP system called *Aleph* [**A** **L**earning **E**ngine for **P**roposing **H**ypotheses], a platform for exploring ILP ideas. All of this work was done in the UoA , at the PRG.

From 2000 onwards PRG member Stephen Moyle began to explore how to use ILP techniques to analyse grammar. Working with Srinivasan, and informed by his Aleph system, Moyle applied the theories of [4] developed in studying natural language to analysing strings of computer code in the same way. In work with John Heasman in 2003, Moyle applied ILP to intrusion detection, the identification of potential breaches in computer security policy. They demonstrated that logic programming was a suitable formalism for specifying the semantics of attacks, and that logic programs could then be used as a means of detecting attacks in previously unseen inputs. The advantage of ILP was that it could be used to induce detection clauses from examples of attacks, and that accurate theories could be generated from very few attack examples. This work was subsequently patented by Heasman and Moyle when setting up their company [5] (and see Section 4).

Heasman and Moyle then tackled the problem of how to automatically create detection rules for a particular type of computer hack known as buffer overflow exploits. By treating the problem of generating buffer overflow strings as a form of grammar, they showed in 2003 that ILP was extremely effective at recovering the attack strategy being used by the attacker in assembling buffer overflow strings [6].

3. References to the research (indicative maximum of six references)

The three asterisked outputs best indicate the quality of the underpinning research.

***[1] Stephen Muggleton, Luc De Raedt: Inductive Logic Programming: Theory and Methods. J. Log. Program. 19/20: 629-679 (1994). DOI: 10.1016/0743-1066(94)90035-3**

This paper outlined the key aspects of the new field, including a 'model theory' for ILP, a generic ILP algorithm and a 'proof theory', as well as applications and implementations of ILP.

*** [2] Stephen Muggleton: Inverse Entailment and Progol. New Generation Comput. 13(3&4): 245-286 (1995). DOI: 10.1007/BF03037227**

This paper describes the implementation of ILP in Prolog.

[3] Stephen Muggleton. Stochastic logic programs. Advances in inductive logic programming 32 (1996): 254-264. *This paper develops ILP by adding probability.*

[4] James Cussens. Part-of-speech tagging using Progol. In Inductive Logic Programming: Proceedings of the 7th International Workshop (ILP-97). LNAI 1297, pp 93-108. Springer, 1997. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.7266

This paper, written when the author was a researcher on the ILP2 project at Oxford, shows the use of ILP to build parts-of-speech taggers from very large example corpuses in natural language applications.

[5] John Heasman and Steve Moyle. Intrusion Detection System. 2004. Patent application no. PCT/GB2003/001466. Publication no. WO2003090046 A2. https://www.google.com/patents/WO2003090046A2?dq=heasman&ei=ixngUaWVF7CT0AWR-YGQBQ&cl=en

*** [6] Steve Moyle, John Heasman. Machine Learning to Detect Intrusion Strategies. KES 2003: 371-378 DOI 10.1007/978-3-540-45224-9_52**

This paper describes how ILP can be used to detect data intrusions.

4. Details of the impact (indicative maximum 750 words)

Route to impact

In 2003, with the assistance of Isis Innovation, Moyle founded SafeTalk Intelligent Technologies, which went on to become Secerno Ltd, a company which used ILP and the techniques described above to protect databases by making policy-based, real-time decisions to prevent unlawful access of data.

It is evident from the patent [5] that was filed by Isis to support this company that the work was based upon Moyle's earlier work on ILP as its description states "The second aspect of the invention preferably uses Inductive Logic Programming techniques (ILP - a form of machine learning), described below, to suggest new intrusion detection rules for inclusion into the system, based on examples of sinister traffic. This not only frees the EDS administrator from having to manually analyse new attacks, but it also means that suggested rules benefit from the semantic matching referred to above."

At Secerno Moyle set out to provide a solution to the problem of database fraud, which is estimated to cost over $2bn per year. Building on his previous work on intrusion detection, Moyle pursued a symbolic machine learning solution to a specific computer security vulnerability known as SQL injections. Such incidents see a database subverted by someone misusing an application trusted by the database, and they continue to account for up to 25% of all data breaches. Moyle developed an approach called grammatical clustering and produced a prototype system that took SQL grammar and examples of the legitimate SQL code sent to a database and extracted a representation of the set of permitted SQL statements. This set was then capable of detecting all anomalous SQL statements formed by a malicious user through SQL injection – and, in turn, blocking them. $3.8m in Series A funding enabled Secerno to create DataWall, first released in

2006, with its technology patented [A]. This second patent is based around stochastic logic programming as introduced in [3] above.

Impact since 2008

In July 2008, $16m of Series B funding was secured as a result of the initial success of DataWall. At this time, DataWall was the only system available that could block with zero false positives, and the product had just been made available as a virtualised appliance on the Vmware platform. The new funding enabled Secerno to improve DataWall and release a series of new versions, including one specifically designed for SMEs in December 2008 [B]. Version 4.1 in January 2010 actively monitored up to 230,000 transactions per second with the ability to block all threats in real-time.

DataWall provided machine-learning technology to observe how applications access specific databases, thereby allowing IT managers to control and protect their data assets from known and unknown, external and internal threats. Unlike conventional approaches of the time, which permitted access to every request except those on a "blacklist", Moyle introduced the first commercial application of a "whitelist" approach where database queries are blocked in real-time unless they are on the approved list. This would be an onerous and unworkable process were the system manager required to develop and maintain complex policy configurations for every scenario, but Moyle's work demonstrated a method of automatically generating a bespoke whitelist of permitted, policy compliant, activities for each database it was attached to. DataWall monitored database activity on the network to prevent unauthorized access, SQL injections, privilege or role escalation, and other external and internal attacks in real time [C], thus supporting organisations to reduce security risks to their data.

As a result of the Series B funding, Secerno was also able to increase its workforce to 55 internal jobs (mostly in the UK) in 2009, and from then recruited over 11 collaborative reseller partnerships. The further development of the product overcame initial industry reluctance to accept whitelisting by demonstrations and pre-sales trials which repeatedly illustrated the superiority of the approach whilst detecting system inadequacies and on-going but previously undetected frauds. Customers included the Skipton Building Society who bought DataWall in November 2008 after losing an encrypted laptop with the personal details of 14,000 customers and receiving a warning from the Information Commissioner's Office [D]. Users of DataWall for SMEs included Let Check, a leading UK tenant referencing company providing instant credit checks and real-time verification on a wide range of personal ID [B].

Impressed by DataWall's ability to block unauthorised activity in real time, in June 2010 Oracle Corporation acquired Secerno for an undisclosed sum in order to obtain DataWall, which was judged to significantly complement and augment Oracle's existing products. A senior vice president of Oracle Database Server Technologies stated that the Secerno acquisition was 'in direct response to increasing customer challenges around mitigating database security risk' and that 'Oracle's complete set of database security solutions and Secerno's technology will provide customers with the ability to safeguard their critical business information.' [E] Renamed Database Firewall, the product now forms part of Oracle's portfolio of database security solutions, and is considered the company's first line of defence for both Oracle and non-Oracle databases. Oracle is unable to divulge internal figures relating to sales, profits and market success. However, in an industry where commercial buy-outs are often a means of eradicating competition, the Secerno acquisition has clearly been a good investment for Oracle. Though there are no publicly available sales figures or financial breakdowns, Oracle continues to sell and develop the Oracle Database Firewall – now combined into Oracle Audit Vault and Database Firewall 12.1.1 – which indicates that it has been a successful purchase for the company.

Customers using Database Firewall are typically large corporations, such as TransUnion – a global financial services company which serves 45,000 industrial and 500 million consumer clients across 32 countries. It uses Oracle Database Firewall as its first line of database security across all of its servers in order to analyse and log all SQL traffic, and in turn is able to develop policies that allow it to block malicious traffic and keep its customers' data safe [F]. This has reduced successful hacking incidents on their wide range of databases. Database Firewall can provide a faster response to automatically detect unauthorised database activities that violate security policies, and thwart perpetrators from covering their tracks [G]. Another large user is T-Mobile, which uses Database Firewall to protect the personal data of their 35 million users [G].

In a February 2013 independent review of Oracle Database Firewall, European IT analysis firm KuppingerCole found that 'the detection of events is based on the grammar-based approach which provides a very high level of accuracy and allows finding a good balance between avoiding false negatives at all and minimizing the number of false positives. Due to the flexible deployment models, support for in-line blocking and monitoring, remote monitoring, out-of-band monitoring, and high availability mode, the solution is well able to support any kind of scenario' [H]. In addition to protecting data from a commercial and customer service perspective, companies are also required to comply with various regulations (such as those introduced by the Sarbanes Oxley Act 2002, the GLBA and the Payment Card Industry Data Security Standard). Database Firewall makes this process far simpler and enables compliance where it may not have previously existed.

5. Sources to corroborate the impact (indicative maximum of 10 references)

[A] Stephen Anthony Moyle. Method, Computer Program and Apparatus for Analysing Symbols in a Computer system. 2009. US Patent Application number: 12/187,104. Publication number: US 2009/0055166 A1 https://www.google.com/patents/US20090055166
Patent taken out by Secerno on Moyle's grammatical clustering approach to intrusion detection, based on stochastic logic programming.
[B] http://www.businessmag.co.uk/News/Technology/Oxford--Secerno-launches-SME-database-security.aspx
Article confirming the launch of DataWall for SMEs and the purchase of this product by Let Check.
[C] http://www.oracle.com/technetwork/products/database-firewall/index.html
Oracle web page describing the product's capabilities.
[D] http://www.itpro.co.uk/608619/skipton-acts-on-ico-warning
Article confirming the purchase of DataWall by Skipton Building Society in November 2008.
[E] http://www.oracle.com/us/corporate/press/075001
Oracle press release confirming their purchase of Secerno and the rationale for it.
[F] http://www.oracle.com/technetwork/issue-archive/2012/12-jul/o42dbsecurity-1652955.html
Article from July 2012 Oracle magazine, confirming Transunion's implementation of Database Firewall.
[G] http://www.oracle.com/us/products/database/security/audit-vault-database-firewall/overview/index.html
Videos on the Oracle website from TransUnion and T-Mobile in which key staff responsible for database security at each company corroborate their use of Database Firewall and outline the advantages of the product to them.
[H] http://www.oracle.com/us/corporate/analystreports/kuppinger-cole-audit-vault-1911963.pdf
February 2013 independent summary report from KuppingerCole outlining the advantages of Database Firewall.